Mulukala, S. K. N., Nishad, R., Kolligundla, L. P., Saleem, M. A., Prabhu, N. P., & Pasupulati, A. K. (2016). *In silico* Structural characterization of podocin and assessment of nephrotic syndrome-associated podocin mutants. *IUBMB Life*, *68*(7), 578-588. https://doi.org/10.1002/iub.1515

In Silico Structural Characterization of Podocin and Assessment of Nephrotic Syndrome Associated Podocin Mutants

Journal:	*IUBMB Life*
Manuscript ID	TBMB-16-0096-WJW.R1
Wiley - Manuscript type:	Research Communication
Date Submitted by the Author:	n/a
Complete List of Authors:	Mulukala, NSK; Biochemistry Nishad, Rajkishor; Biochemistry Kolligundla, Lakshmi Prasanna; University of Hyderabad School of Life Sciences, Biochemistry Saleem, Moin; University of Bristol Prabhu, N Prakash; University of Hyderabad, Department of Biotechnology & Bioinformatics Pasupulati, Anil Kumar; Biochemistry
Keywords:	

In Silico Structural Characterization of Podocin and Assessment of Nephrotic Syndrome Associated Podocin Mutants

NSK. Mulukala[1], R. Nishad[1], LP. Kolligundla[1], Moin A. Saleem[2], N. Prakash Prabhu[3], Anil Kumar Pasupulati[#1]

1-Department of Biochemistry, University of Hyderabad, India; 2- Academic Renal Unit University of Bristol, Bristol, UK; 3- Department of Biotechnology & Bioinformatics, University of Hyderabad, India;

Running title: Podocin structure prediction & mutation analysis

Keywords: Nephrotic syndrome, Proteinuria, Podocytes, Podocin, Slit-diaphragm, Molecular modelling.

#Correspondence to:
Anil Kumar Pasupulati
Department of Biochemistry,
School of Life Sciences
University of Hyderabad,
Gachibowli, Hyderabad, India – 500046
E-mail: pasupulati.anilkumar@gmail.com
Phone: 91-40-23134519

Abstract

Nephrotic syndrome (NS) is manifested by hyperproteinuria, hypoalbuminemia and edema. *NPHS2* that encodes podocin was found to have most mutations among the genes that are involved in the pathophysiology of NS. Podocin, an integral membrane protein belonging to stomatin family, is expressed exclusively in podocytes and is localized to slit-diaphragm (SD). Mutations in podocin are known to be associated with steroid resistant nephrotic syndrome (SRNS) and rapid progression to end-stage renal disease (ESRD), thus signifying its role in maintaining SD integrity and podocyte function. The structural insights of podocin are not known and the precise mechanism by which podocin contributes to the architecture of SD is yet to be elucidated. In this study we deduced a model for human podocin, discussed the details of trans-membrane localisation, intrinsically unstructured regions and provide an understanding of how podocin interacts with other SD components. Intra-protein interactions were assessed in wild type podocin and in some of its mutants that are associated with idiopathic NS. Mutations in podocin alter the innate intra-protein interactions affecting the native structure of podocin and its ability to form critical complex with sub-podocyte proteins.

Introduction:

The kidneys are vital organs that help maintain body homeostasis by regulating blood pressure, acid-base, electrolyte and water balance. Human kidney constitutes a million nephrons that collectively perform three key events including a) glomerular filtration of water and small molecules from renal plasma; b) tubular reabsorption of glomerular filtrate; c) tubular secretion of metabolic waste products into the filtrate. Thus, glomerulus in concert with tubular region of the nephron tightly regulates the composition of glomerular filtrate and ensures almost protein-free ultra-filtrated urine. Glomerulus, where initiation of filtration occurs contains a tuft of capillaries and several resident cell types that include mesangial cells, endothelial cells and glomerular visceral epithelial cells, also known as podocytes.

Proteinuria is a hallmark of renal damage in several glomerular diseases due to the alterations in glomerular filtration barrier (GFB) (1, 2). The three components that constitute GFB include fenestrated glomerular endothelial cells, glomerular basement membrane (GBM) and glomerular visceral epithelial cells, known as podocytes. A wealth of literature highlighted that podocytes are critical for glomerular filtration (3, 4). Podocytes are terminally differentiated epithelial cells with large cell body consisting of organelles. These cells possess a main cell body, major foot process made of microtubules and intermediate filaments and numerous secondary foot process that play a predominant role in attaching podocytes to the GBM. The secondary foot process of neighboring podocytes forms an adherent junction called slit-diaphragm (SD) that forms a sole contact between neighboring podocytes and ensures size, shape and charge selective permeability to the GFB (1). SD is made up of proteins like nephrin, podocin, CD2-associated protein (CD2AP), ZO-1, and P-cadherin. Mutations in the SD proteins alter podocyte architecture and permselectivity, which eventually abate the function of GFB culminating in proteinuria (5). Mutations in podocyte proteins could target the function of podocyte by affecting the structure of the SD, by perturbing the intricate podocyte cytoskeleton, by breaking cell-matrix interactions and by blocking important signaling pathways. All these events manifest in the effacement of podocyte foot processes, disruption of GFB and proteinuria (6).

Nephrotic syndrome (NS) is a nonspecific renal disorder characterized by heavy proteinuria, hypoalbuminemia, and edema. Idiopathic nephrotic syndrome (INS) that occur most frequently in children, represent >90 % of NS cases between 1 and 10 years of age and 50% of NS cases post 10 years of age (7). Electron microscopy of renal biopsies from INS subjects revealed diffuse foot process effacement, minimal change disease (MCD) and focal segmental glomerulosclerosis (FSGS). INS that occurs mainly in children is classified into steroid-sensitive (SSNS) and steroid-resistant NS (SRNS), depending on their response to corticosteroid therapy. SRNS has a poor prognosis and often leads to end-stage renal disease (ESRD). Mutations in more than 20 genes that predominantly encode podocyte proteins have been identified in monogenic forms of SRNS (8). Among those genes, *NPHS2*, encoding podocin, is the most frequently mutated gene responsible for up to 18% of total SRNS cases (9, 10). Podocin consists of 383 amino acids and is expressed exclusively in podocytes, where in it localizes to SD as an oligomer (10-12). Podocin belongs to stomatin family and is hypothesized to form hairpin-like intramembrane loop and intracellular C-terminus (12). It was reported that

podocin interacts with other proteins of SD such as CD2AP and nephrin with its C-terminus (13). The interactions between podocin and other proteins are considered to be critical for the maintenance of intact SD architecture. Mutation in podocin lead to the early onset of SRNS typically before 6 years of age and rapidly progresses to ESRD within a decade, indicating the instrumental role of podocin in podocyte biology and in maintaining the integrity of GFB.

Podocin mutants such as R138Q and V180M were observed to be restricted to endoplasmic reticulum, intracellular vesicles or form inclusion bodies thus preventing its localization to SD (14). However, it is yet to be known how mutations in podocin are associated with morphological changes ranging from MCD and FSGS that manifest in heavy proteinuria in SRNS. Further studies are warranted to understand how mutations in podocin will distort the innate interactions with its sub SD proteins. In this study we attempted to generate 3D models for podocin and its mutants that are associated with SRNS to gain insights about how mutations affect the protein structure and contribute towards altered SD architecture and podocyte function.

Methodology:

BLAST analysis of podocin sequence and sequence alignment: Human podocin sequence (accession ID: Q9NP85) was obtained from ExPASy database. Since crystal structure for podocin is not available, a reliable model was built based on sequence similarity search using protein BLAST tool against predetermined protein structures deposited in protein data bank (PDB). A comparative sequence alignment of the identified homologous sequences was done using the tools ClustalW and ClustalX.

Structure prediction using I-TASSER server: 3D structure prediction for podocin sequence bits (1-164 and 362-383) that showed less (<25%) homology was performed using the I-TASSER web server (http://zhanglab.ccmb.med.umich.edu/I-TASSER/). I-TASSER, an *ab initio* based protein structure prediction software works on sequence-to-structure-to-function pattern (15). The server generated five models, which are ranked based on the structure density of the SPICKER clustering. A confidence score (C-score) was calculated based on the statistical significance of the threading profile-profile alignment (PPA) as well as structure convergence of the Monte Carlo simulations. This C-score denotes the quality of the models generated by the server.

Model prediction and stereochemistry analysis of the model: The predicted structure from I-TASSER server and the homologue sequences obtained from the BLAST search were taken as templates to generate a model for podocin. A multiple sequence based modeling was done using Modeller 9.15 software that generates models by satisfying spatial restrains via automated comparative modeling (16). The multiple sequence alignment file in Modeller 9.15 was manually tweaked to match the query and template sequences. The generated model was then analyzed in detail for its protein structure stereochemistry using protein parameters analysis tool known as PROCHECK.

Predicting intrinsically unstructured, trans-membrane regions and assessing intra-protein contacts: Due to the anomalous N- and C-terminal random coils observed in the generated model, a secondary structure analysis was done using PSI-PRED server (http://bioinf.cs.ucl.ac.uk/psipred/). This server uses a two-stage neural network to predict the secondary structure using data generated by Position-Specific Iterated BLAST (PSI-BLAST) (17). Further, the server also extrapolated intrinsically unstructured regions (IURs) (DISOPRED3), trans-membrane (TM) topology prediction (MEMSAT3) and trans-membrane helix prediction (MEMSATSVM) (18-20). However, in order to gain further evidence on IURs in podocin, the sequence was trained exhaustively in various Critical Assessment of Protein Structure Prediction (CASP) validated servers that are: Genesilico meta-disorder service (Metadisorder, MetadisorderMD, MetadisorderMD2, Metadisorder3D) (21), PONDR (22), DisProt (23), IUPred (24), Dis-EMBL (disorder by loops/coils definition, disorder by HOT-loops, Remark-465) (25), SPINE-D (26), MFDp (DISOclust, DISOpred, IUPredL, IUPredS) (27) and PredictProtein (28). The PSI-PRED server and the PredictProtein server extrapolated TM helices in addition to predicting IURs (29). The intra-protein interactions such as main chain-main chain, main chain-side chain, side chain-side chain, hydrophobic interactions for a given set of 3D coordinates of the protein was assessed using protein interactions calculator (PIC) server (30).

Building models for podocin mutants: We have selected 6 mutants of podocin (R3G, P89T, R322Q, R322P, H325Y and V370G) that were shown to associate with SRNS from HGMD database (31, 32). The criteria for selecting these particular mutants was they occur in predicted IURs as described above. Models for these mutants were built using Modeller 9.15 software considering wild type podocin as a template. Further, stereochemistry of mutant models was analysed using PROCHECK software.

Results:

Sequence alignment and homology modeling of Podocin: Podocin shows highest homology with stomatin from *H. sapiens* (47% identity and 67% similarity) and with mechanosensory protein-2 (MEC-2) from *C. elegans* (44% identity and 65% similarity) (10). Considering this homology as a base, a protein BLAST (pBLAST) search was performed with human podocin sequence against PDB database. The search yielded two prominent templates: 1. Chain A, crystal structure of a core domain of stomatin from *Pyrococcus horikoshii* (PDB ID: 3BK6) (33) and 2. Chain C, RecBCD: DNA complex from *Escherichia coli* (PDB ID: 1W36) (34). These two templates covered the podocin sequence from 165 to 361 amino acids.

Although, two homologous templates were identified from the pBLAST search, the N- and C-terminal regions (1-164 and 362-383 aminoacids) of podocin showed < 25%homology to build a reliable full-length model. Hence, an *ab initio* structure prediction was done using I-TASSER server for these regions of podocin. The templates thus obtained from the BLAST search and I-TASSER prediction were used to build a multiple sequences based model (Fig 1). The Ramachandran plot for podocin model revealed 85.5% of amino acids in most favourable regions; 10.9% amino acids in additionally allowed regions; 2.7% of amino acids in generously

allowed regions and 0.9% amino acids in disallowed regions (Fig 2A). It is noteworthy that the prediction revealed the presence of random coils at both N- and C-terminals (Fig 2B and C). We speculated that these random coils could be IURs, since BLAST search with N-terminus podocin sequence showed homology with Influenza A virus nucleoprotein that showed unstructured regions. Therefore, we further continued our quest of gaining insights on podocin structure by analyzing whether these coils are intrinsically unstructured. Since IURs form anchors and signalling motifs, we speculate that IURs in podocin may involve in the formation of scaffolding complex with CD2AP and TRPC6 ion channel besides facilitating signalling via nephrin (13, 35).

Predicting intrinsically unstructured and transmembrane regions: Podocin sequence was screened employing seven standalone and three meta IUR prediction servers, which predicted an N-terminal IUR (1 to 98 residues) and C-terminal IUR (347 to 383 residues) (Table 1 & Fig 3). Further, we have employed Genesilico meta-disorder prediction server to assess the presence of IURs, if any in other stomatin family proteins. Interestingly, all stomatin family proteins showed IURs at various regions in their respective sequences (Table 2). These results strongly suggest the presence of IURs at both N- and C-terminus of podocin. The PSI-PRED server identified the presence of protein binding regions between residues 1-7, 63-69, 77-89, 353-356, 363 and 367-377 in the identified IURs, which perhaps explains scaffolding complex formation of podocin with CD2AP, nephrin (13) and extended complex with TRPC6 (35). The server also predicted a transmembrane (TM) region in podocin spanning from ~101 to 125 amino acids.

It has been reported that stomatin family members exhibit cytoplasmic C-and N-terminus (36), thus podocin is also expected to show similar structural properties to that of other stomatin family members. Contrastingly, PSI-PRED analysis showed two possible orientations for N- and C-terminals of podocin. The prediction suggests that N-terminus is cytoplasmic and C-terminus is extracellular, if the TM region is restricted between 98 to 125 residues (Fig 4A). Alternatively, it was also predicted that N-terminus is extracellular and C-terminus is cytoplasmic, if the TM region is localized between 103 to 127 residues (Fig 4B).

Mutations in podocin alter intra-protein interactions: To analyse how mutations in podocin, specifically in the predicted IURs alter its structure and perturb its interactions with sub-podocyte proteins, six podocin mutants were selected and models for these mutants were generated using wild type podocin model as a template. The Ramachandran plots for the mutants showed minimal variation with respect to that of wild type podocin despite the fact that these podocin mutants are associated with nephrotic syndrome (Table 3).

The mutants were then analysed to identify the changes if any, in the intra-protein interactions with the help of PIC server. This server calculates various interactions based on the given coordinates of 3D structure of a protein (30). Substantial differences in ‘main chain-main chain’, ‘main chain-side chain’, ‘side chain-side chain’ and ‘hydrophobic interactions’ were observed when compared to wild type podocin.

These altered intra-protein interactions in podocin mutants showed considerable distortions in their secondary structures. All podocin mutants we analysed showed the formation of an

extended alpha helix between residues 307-325, except in mutant R3G wherein, T315 residue was found to be a part of the alpha helix (residues 316-322) that was observed in wild type protein (Fig 5). Two β-sheets (170-181 & 185-197 residues) in the wild type podocin were trimmed (170-180 and 186-197 residues) in the case of H325Y mutant. Furthermore, in mutants R3G, P89T, R322Q and V370G, stable alpha helices (residues 151-156) were replaced with 3_{10} helices (residues 153-155) (Fig 5).

Root mean square deviations (RMSD) of podocin mutants with that of podocin wild type showed that R3G mutant deviate by 1.23 Å when compared to wild type podocin. Similarly P89T, R322Q, R322P, H325Y and V370G mutants showed 1.01 Å, 1.04 Å, 1.12 Å, 0.68 Å and 0.55 Å deviations when compared to the wild type podocin (Fig 6).

Discussion:

In this study, we have provided insights of podocin structure that was generated using a multiple template based structure prediction analysis. This study predicted the orientation of transmembrane (TM) region and presence of random coils, which we have identified as IURs at N-and C-terminal regions of podocin. We have also identified protein binding regions in these IURs, which could play a critical role in complex formation with sub-podocyte SD proteins. Intra-protein interactions among various amino acid residues were assessed in wild type and in some of podocin mutants that were associated with steroid resistant form of idiopathic NS. Mutations in podocin alter innate intra-proteins interactions affecting the native structure of podocin and its ability to form critical complex with sub-podocyte proteins.

The structure of human podocin was designed based on multiple sequence homology modelling and the model thus generated was evaluated for protein stereochemistry using PROCHECK software. Interestingly, it was observed that both N- and C-terminal regions of podocin showed random coils. pBLAST search with the N-terminal region of podocin (1-156) showed homology with the nucleoprotein of influenza A virus (PDB ID: 3ZDP - identity: 33%, query coverage: 39%), which displays a flexible disordered region (402-428 residues). This region was identified to connect the tail loop and main body of the nucleoprotein that points towards the RNA binding surface (37). Therefore, we have speculated the presence of intrinsically unstructured regions (IURs) in the podocin.

IURs do not form a fixed 3D structure under physiological conditions either in their entireties or they may contain intrinsically disordered regions. They resemble the denatured states of ordered proteins and are described as an ensemble of rapidly inter-converting alternative structures (38). IURs take up different structures upon binding to different targets, and thereby exhibit functional flexibility through the formation of fuzzy complexes (39). Comprehensive analysis by various CASP validated servers revealed that podocin consists of both N-terminal (1 to 98 amino acids) and C-terminal (347 to 383 amino acids) IURs. Our predictions also identified protein binding regions in these IURs which makes us speculate that these IURs may help podocin to serve as an anchor, signalling motif or help in interacting with neighbouring proteins to form a complex, thus maintaining podocyte SD integrity (13, 35, 40).

However, a detailed elucidation of how nephrin-CD2AP-podocin complex is formed is greatly warranted, which could be accomplished by solving the high-resolution crystal structure of these protein(s) complexes.

While we were analysing IURs in podocin, the PSI-PRED server predicted the presence of a TM region (100-125 amino acids) in podocin, which correlated with results from an earlier study (10). Interestingly, TM region that was identified in podocin is homologous to that present in other stomatin family members (10). It was not known whether N-terminal region is cytoplasmic or extracellular, however, it was reported that C-terminus is cytoplasmic (13). Nevertheless, the possibility of podocin to form a hairpin loop to interact with membrane from cytoplasmic side cannot be ruled out as other stomatin family proteins possess cytoplasmic C-terminus and forms hairpin loops (10). Further, it was shown that N- and C-terminal specific podocin antibodies co-localized at the cytoplasmic face of the plasma membrane suggesting presence of hairpin loop structure (41).

Mutations in *NPHS2* cause early onset of SRNS and ensue rapid progression to ESRD. We have selected 6 mutations from HDMG database and generated models for them. We attempted to understand how mutations specifically in the intrinsically unstructured regions affected the macromolecular assembly. It was found that mutations elicit significant changes in intra-protein interactions such as main chain-main chain, main chain-side chain, hydrophobic and side chain-side chain contacts. These changes showed considerable differences in protein secondary structure. It could be speculated that alterations in intra-protein interactions bring about changes in microenvironment and compromise stereochemistry and macromolecular assembly of podocin with its sub-podocyte components. Further, the structural changes observed at distant regions of mutations may be viewed along with the larger changes observed in the intermittent residues as well. For instance, R3G mutant shows change of secondary structure from alpha-helix to 3_{10}-helix in the region of 151-156, thus bringing about distant structural changes via the changes in intermittent residues, which could be observed in other mutant forms as well. This could possibly explain how mutations in podocin cause detrimental changes in a larger scale to the podocyte SD structure and loss of podocyte function.

It is noteworthy that podocin mutants that were selected in this study such as R3G, R322Q, R322P, H325Y and V370G were manifested in FSGS and the mutation P89T resulted in MCD (31, 32, 42). It was shown that podocin mutant H325Y progresses to ESRD (43). From the observed data we speculated that even the slightest variation in the conformation of the protein secondary structure could alter the macromolecular assembly of podocin with sub-SD proteins due to alterations of podocin backbone, which in turn could eventually manifest in altered architecture of SD and proteinuria in NS.

Acknowledgements: We acknowledge Department of Science and Technology, India for providing research grants IFA11-LSBM-02 and SB/FT/LS-307/2012 (to PAK). We acknowledge the resources provided by biotechnology infrastructure facility (BIF) of the University of Hyderabad.

Conflict of Interest: Authors do not have any conflict of interest to declare.

References

1. Anil Kumar, P., G.I. Welsh, M.A. Saleem, and R.K. Menon, *Molecular and cellular events mediating glomerular podocyte dysfunction and depletion in diabetes mellitus.* Front Endocrinol (Lausanne), 2014. **5**: p. 151.
2. Bazzi, C., O. Bakoush, and L. Gesualdo, *Proteinuria: from molecular to clinical applications in glomerulonephritis.* Int J Nephrol, 2012. **2012**: p. 424968.
3. Kumar, P.A., C.B. Frank, and K.M. Ram, *The Glomerular Podocyte as a Target of Growth Hormone Action: Implications for the Pathogenesis of Diabetic Nephropathy.* Current Diabetes Reviews, 2011. **7**(1): p. 50-55.
4. Greka, A. and P. Mundel, *Cell biology and pathology of podocytes.* Annu Rev Physiol, 2012. **74**: p. 299-323.
5. Somlo, S. and P. Mundel, *Getting a foothold in nephrotic syndrome.* Nat Genet, 2000. **24**(4): p. 333-5.
6. Gigante, M., M. Piemontese, L. Gesualdo, A. Iolascon, and F. Aucella, *Molecular and genetic basis of inherited nephrotic syndrome.* Int J Nephrol, 2011. **2011**: p. 792195.
7. *Nephrotic syndrome in children: Prediction of histopathology from clinical and laboratory characteristics at time of diagnosis.* Kidney Int, 1978. **13**(2): p. 159-165.
8. Sadowski, C.E., S. Lovric, S. Ashraf, W.L. Pabst, H.Y. Gee, et al., *A single-gene cause in 29.5% of cases of steroid-resistant nephrotic syndrome.* J Am Soc Nephrol, 2015. **26**(6): p. 1279-89.
9. Benoit, G., E. Machuca, and C. Antignac, *Hereditary nephrotic syndrome: a systematic approach for genetic testing and a review of associated podocyte gene mutations.* Pediatr Nephrol, 2010. **25**(9): p. 1621-32.
10. Boute, N., O. Gribouval, S. Roselli, F. Benessy, H. Lee, et al., *NPHS2, encoding the glomerular protein podocin, is mutated in autosomal recessive steroid-resistant nephrotic syndrome.* Nat Genet, 2000. **24**(4): p. 349-354.
11. Miner, J.H, *Focusing on the glomerular slit-diaphragm: Podocin enters the picture. Am J Pathol,* 2002. 3-5.
12. Schurek, E.M., Völker, L.A., Tax, J., Lamkemeyer, T,. Rinschen, M.M, et al., *A disease-causing mutation illuminates the protein membrane topology of the kidney-expressed prohibitin homology (PHB) domain protein podocin*. J Biol Chem, 2014. **289**(16):p. 11262-71.
13. Schwarz, K., M. Simons, J. Reiser, M.A. Saleem, C. Faul, et al., *Podocin, a raft-associated component of the glomerular slit diaphragm, interacts with CD2AP and nephrin.* J Clin Invest, 2001. **108**(11): p. 1621-9.
14. Nishibori, Y., L. Liu, M. Hosoyamada, H. Endou, A. Kudo, et al., *Disease-causing missense mutations in NPHS2 gene alter normal nephrin trafficking to the plasma membrane.* Kidney Int, 2004. **66**(5): p. 1755-65.
15. Zhang, Y., *I-TASSER server for protein 3D structure prediction.* BMC Bioinformatics, 2008. **9**: p. 40.
16. Sali, A. and T.L. Blundell, *Comparative protein modelling by satisfaction of spatial restraints.* J Mol Biol, 1993. **234**(3): p. 779-815.
17. Jones, D.T., *Protein secondary structure prediction based on position-specific scoring matrices.* J Mol Biol, 1999. **292**(2): p. 195-202.

18. Jones, D.T. and D. Cozzetto, *DISOPRED3: precise disordered region predictions with annotated protein-binding activity.* Bioinformatics, 2015. **31**(6): p. 857-63.
19. Jones, D.T., W.R. Taylor, and J.M. Thornton, *A model recognition approach to the prediction of all-helical membrane protein structure and topology.* Biochemistry, 1994. **33**(10): p. 3038-49.
20. Nugent, T. and D.T. Jones, *Transmembrane protein topology prediction using support vector machines.* BMC Bioinformatics, 2009. **10**: p. 159.
21. Kozlowski, L.P. and J.M. Bujnicki, *MetaDisorder: a meta-server for the prediction of intrinsic disorder in proteins.* BMC Bioinformatics, 2012. **13**(1): p. 1-11.
22. Obradovic, Z., K. Peng, S. Vucetic, P. Radivojac, C.J. Brown, et al., *Predicting intrinsic disorder from amino acid sequence.* Proteins, 2003. **53 Suppl 6**: p. 566-72.
23. Peng, K., S. Vucetic, P. Radivojac, C.J. Brown, A.K. Dunker, et al., *Optimizing long intrinsic disorder predictors with protein evolutionary information.* J Bioinform Comput Biol, 2005. **3**(1): p. 35-60.
24. Dosztanyi, Z., V. Csizmok, P. Tompa, and I. Simon, *IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content.* Bioinformatics, 2005. **21**(16): p. 3433-4.
25. Linding, R., L.J. Jensen, F. Diella, P. Bork, T.J. Gibson, et al., *Protein disorder prediction: implications for structural proteomics.* Structure, 2003. **11**(11): p. 1453-9.
26. Zhang, T., E. Faraggi, B. Xue, A.K. Dunker, V.N. Uversky, et al., *SPINE-D: accurate prediction of short and long disordered regions by a single neural-network based method.* J Biomol Struct Dyn, 2012. **29**(4): p. 799-813.
27. Mizianty, M.J., W. Stach, K. Chen, K.D. Kedarisetti, F.M. Disfani, et al., *Improved sequence-based prediction of disordered regions with multilayer fusion of multiple information sources.* Bioinformatics, 2010. **26**(18): p. i489-96.
28. Schlessinger, A., M. Punta, G. Yachdav, L. Kajan, and B. Rost, *Improved disorder prediction by combination of orthogonal approaches.* PLoS One, 2009. **4**(2): p. e4433.
29. Rost, B., P. Fariselli, and R. Casadio, *Topology prediction for helical transmembrane proteins at 86% accuracy.* Protein Sci, 1996. **5**(8): p. 1704-18.
30. Tina, K.G., R. Bhadra, and N. Srinivasan, *PIC: Protein Interactions Calculator.* Nucleic Acids Res, 2007. **35**(Web Server issue): p. W473-6.
31. Berdeli, A., S. Mir, O. Yavascan, E. Serdaroglu, M. Bak, et al., *NPHS2 (podicin) mutations in Turkish children with idiopathic nephrotic syndrome.* Pediatr Nephrol, 2007. **22**(12): p. 2031-40.
32. Caridi, G., M. Gigante, P. Ravani, A. Trivelli, G. Barbano, et al., *Clinical features and long-term outcome of nephrotic syndrome associated with heterozygous NPHS1 and NPHS2 mutations.* Clin J Am Soc Nephrol, 2009. **4**(6): p. 1065-72.
33. Yokoyama, H., S. Fujii, and I. Matsui, *Crystal structure of a core domain of stomatin from Pyrococcus horikoshii Illustrates a novel trimeric and coiled-coil fold.* J Mol Biol, 2008. **376**(3): p. 868-78.
34. Singleton, M.R., M.S. Dillingham, M. Gaudier, S.C. Kowalczykowski, and D.B. Wigley, *Crystal structure of RecBCD enzyme reveals a machine for processing DNA breaks.* Nature, 2004. **432**(7014): p. 187-93.
35. Huber, T.B., B. Schermer, R.U. Muller, M. Hohne, M. Bartram, et al., *Podocin and MEC-2 bind cholesterol to regulate the activity of associated ion channels.* Proc Natl Acad Sci U S A, 2006. **103**(46): p. 17079-86.
36. Lapatsina, L., J. Brand, K. Poole, O. Daumke, and G.R. Lewin, *Stomatin-domain proteins.* Eur J Cell Biol, 2012. **91**(4): p. 240-5.
37. Chenavas, S., L.F. Estrozi, A. Slama-Schwok, B. Delmas, C. Di Primo, et al., *Monomeric nucleoprotein of influenza A virus.* PLoS Pathog, 2013. **9**(3): p. e1003275.

38. Sickmeier, M., J.A. Hamilton, T. LeGall, V. Vacic, M.S. Cortese, et al., *DisProt: the Database of Disordered Proteins.* Nucleic Acids Res, 2007. **35**(Database issue): p. D786-93.
39. Tompa, P. and M. Fuxreiter, *Fuzzy complexes: polymorphism and structural disorder in protein-protein interactions.* Trends Biochem Sci, 2008. **33**(1): p. 2-8.
40. Huber, T.B., M. Kottgen, B. Schilling, G. Walz, and T. Benzing, *Interaction with podocin facilitates nephrin signaling.* J Biol Chem, 2001. **276**(45): p. 41543-6.
41. Roselli, S., O. Gribouval, N. Boute, M. Sich, F. Benessy, et al., *Podocin localizes in the kidney to the slit diaphragm area.* Am J Pathol, 2002. **160**(1): p. 131-9.
42. Sharma. S., Kabra. M, Hari. P, Dinda. A.K, and Bagga. A, *NPHS2 and WT1 mutations in Indian children with steroid-resistant nephrotic syndrome.* Genomic Med, 2008. **2**: p. 241–252
43. Caridi, G., Bertelli. R, Scolari. F, Sanna-Cherchi. S, Di Duca, M, et al., *Podocin mutations in sporadic focal-segmental glomerulosclerosis occurring in adulthood.* Kidney International, 2003. 64 p. 365–8.

Figure Legends:

Figure 1: Templates used for generating full-length podocin model: (A) BLAST alignment of podocin with Chain A, crystal structure of a core domain of stomatin from *Pyrococcus horikoshii* (PDB ID: 3BK6); (B) BLAST alignment of podocin with Chain C, RecBCD: DNA Complex [*Escherichia coli*] (PDB ID: 1W36). (C) Templates used to generate full-length 3D model for podocin. The arrows in red indicate the parts of sequence for which 3D models were predicted using I-TASSER.

Figure 2: Wild type podocin model: A) Ramachandran plot for the wildtype podocin, B) Secondary structure of the generated podocin model, C) modelled podocin structure based on a multiple sequence based homology modelling using templates from pBLAST search and I-TASSER predictions.

Figure 3: Localization of Intrinsically unstructured regions (IURs) in podocin: A consensus data from different IUR prediction servers identified IURs in podocin spanning 1-98 and 347-383 at the N- and C-terminus respectively.

Figure 4: Transmembrane (TM) region predictions for podocin: PSI-PRED predicted two possible conformations for podocin transmembrane region. (A) N-terminus is cytoplasmic and C-terminus is extracellular, if TM region is located between 98-125 residues, and (B) N-terminus is extracellular and C-terminus is cytoplasmic, if the TM region span between 103-127 residues.

Figure 5: Mutations in podocin distorted their secondary structure**:** An extended alpha helix between residues 307-325 was noticed in all the mutants apart from R3G mutant wherein, T315 residue was found to be a part of the alpha helix (residues 316-322) when compared to WT. In mutants R3G, P89T, R322Q and V370G stable alpha helices (residues 151-156) were replaced with 3_{10} helices (residues 153-155). However, in mutation H325Y additional structural changes such as trimming of two β sheets to residues 170-180 and 186-197 was also observed.

Figure 6: Structural alignment of podocin and its mutants: Each figure (A-F) represents structural alignment of wild type podocin model with a mutant. The ball and stick representation in each figure (A: WT vs R3G; B: WT vs P89T; C: WT vs R322P; D: WT vs R322Q; E: WT vs H325Y; F: WT vs V370G) shows the amino acid that was mutated. RMSD for each mutant was calculated using Pymol.

A) Chain A, Crystal Structure Of A Core Domain Of Stomatin From Pyrococcus Horikoshii
Sequence ID: **pdb|3BK6|A** Length: 188 Number of Matches: 1

Score	Expect	Method	Identities	Positives	Gaps	Frame
112 bits(280)	8e-29()	Compositional matrix adjust.	67/172(39%)	108/172(62%)	0/172(0%)	

Features:

```
Query  165  VDLRLQTLEIPFHEIVTKDMFIMEIDAICYYRMENASLLLSSLAHVSKAVQFLVQTTMKR  224
            VDLR Q L++P  E +TKD   + ++A+ Y+R+ +    ++ + +    A   + QTT++
Sbjct  9    VDLRTQVLDVPVQETITKDNVPVRVNAVVYFRVVDPVKAVTQVKNYIMATSQISQTTLRS  68

Query  225  LLAHRSLTEILLERKSIAQDAKVALDSVTCIWGIKVERIEIKDVRLPAGLQHSLAVEAEA  284
            ++    L E+L ER  +    +  +D  T  WGIKV  +EIKDV LPAG+Q ++A +AEA
Sbjct  69   VIGQAHLDELLSERDKLNMQLQRIIDEATDPWGIKVTAVEIKDVELPAGMQKAMARQAEA  128

Query  285  QRQAKVRMIAAEAEKAASESLRMAAEILSGTPAAVQLRYLHTLQSLSTEKPS  336
            +R+ + R+  AEAE+ A+E LR AAEI+S  P A+QLR L T+  ++ +K +
Sbjct  129  ERERRARITLAEAERQAAEKLREAAEIISEHPMALQLRTLQTISDVAGDKSN  180
```

B) Chain C, Recbcd:dna Complex
Sequence ID: **pdb|1W36|C** Length: 1122 Number of Matches: 1

Score	Expect	Method	Identities	Positives	Gaps	Frame
28.9 bits(63)	9.4()	Compositional matrix adjust.	29/116(25%)	53/116(45%)	20/116(17%)	

Features:

```
Query  256  WGIKVERIEIKDVRLPAGLQHSLAVEAEAQRQAKVRMIAAEAEKAA----------SESL  305
            WGI  + +  +++ LPA  QH+        R    RM+   A ++A           FS
Sbjct  504  WGIDDDNV--RELELPATGQHTW-------RFGLTRMLLGYAMESAQGEWQSVLPYDESS  554

Query  306  RMAAEILSGTPAAVQLRYLHTLQSLSTEKPSTVVLPLPFDLLNCLSSPSNRTQGSL  361
             + AE++ G  A++ ++     + L+ E+P    LP+  D+LN    P   T+ ++
Sbjct  555  GLIAELV-GHLASLLMQLNIWRRGLAQERPLEEWLPVCRDMLNAFFLPDAETEAAM  609
```

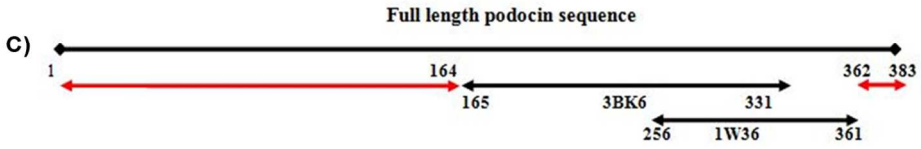

Figure 1: Templates used for generating full-length podocin model: (A) BLAST alignment of podocin with Chain A, crystal structure of a core domain of stomatin from Pyrococcus horikoshii (PDB ID: 3BK6); (B) BLAST alignment of podocin with Chain C, RecBCD: DNA Complex [Escherichia coli] (PDB ID: 1W36). (C) Templates used to generate full-length 3D model for podocin. The arrows in red indicate the parts of sequence for which 3D models were predicted using I-TASSER.
119x141mm (300 x 300 DPI)

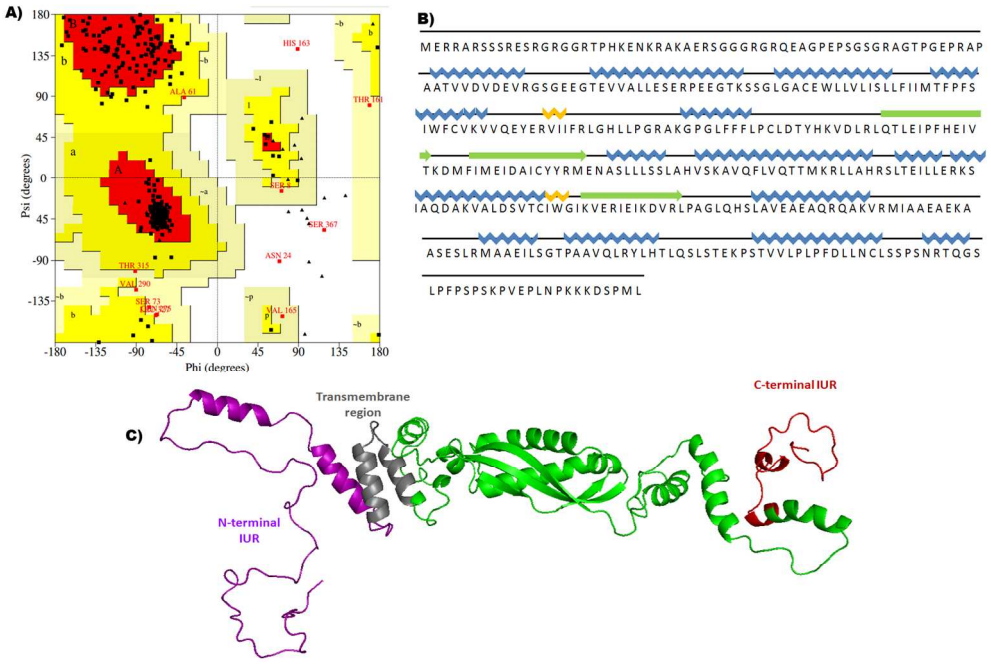

Figure 2: Wild type podocin model: A) Ramachandran plot for the wildtype podocin, B) Secondary structure of the generated podocin model, C) modelled podocin structure based on a multiple sequence based homology modelling using templates from pBLAST search and I-TASSER predictions.
160x108mm (300 x 300 DPI)

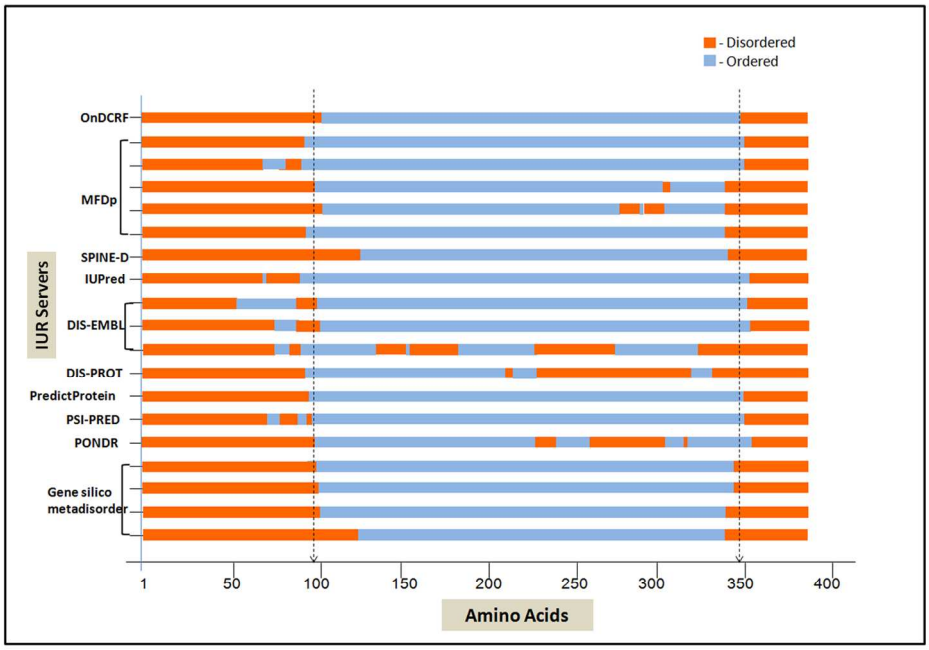

Figure 3: Localization of Intrinsically unstructured regions (IURs) in podocin: A consensus data from different IUR prediction servers identified IURs in podocin spanning 1-98 and 347-383 at the N- and C-terminus respectively.
143x102mm (300 x 300 DPI)

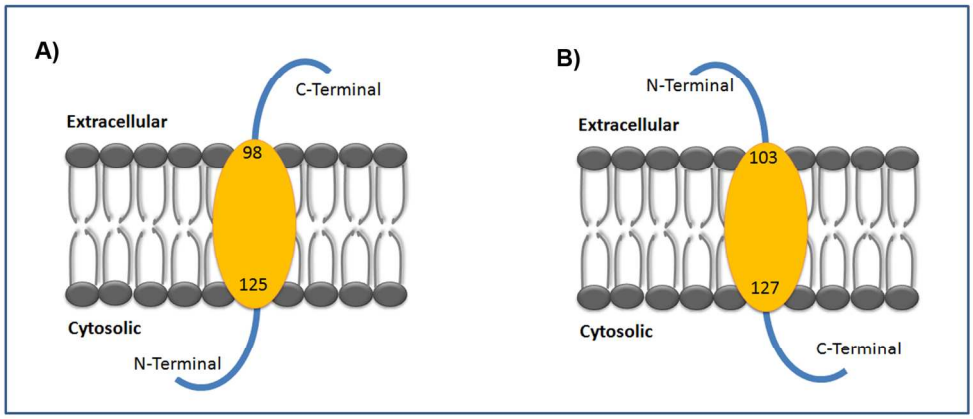

Figure 4: Transmembrane (TM) region predictions for podocin: PSI-PRED predicted two possible conformations for podocin transmembrane region. (A) N-terminus is cytoplasmic and C-terminus is extracellular, if TM region is located between 98-125 residues, and (B) N-terminus is extracellular and C-terminus is cytoplasmic, if the TM region span between 103-127 residues.
149x67mm (300 x 300 DPI)

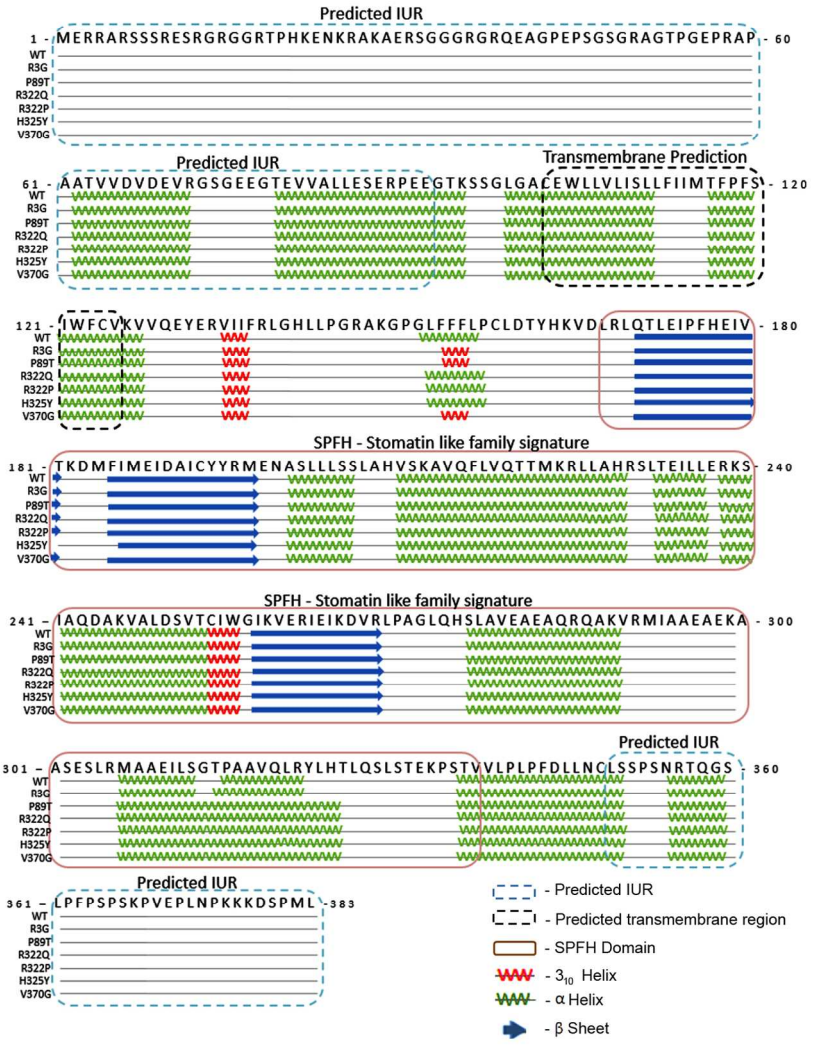

Figure 5: Mutations in podocin distorted their secondary structure: An extended alpha helix between residues 307-325 was noticed in all the mutants apart from R3G mutant wherein, T315 residue was found to be a part of the alpha helix (residues 316-322) when compared to WT. In mutants R3G, P89T, R322Q and V370G stable alpha helices (residues 151-156) were replaced with 310 helices (residues 153-155). However, in mutation H325Y additional structural changes such as trimming of two β sheets to residues 170-180 and 186-197 was also observed.
119x160mm (300 x 300 DPI)

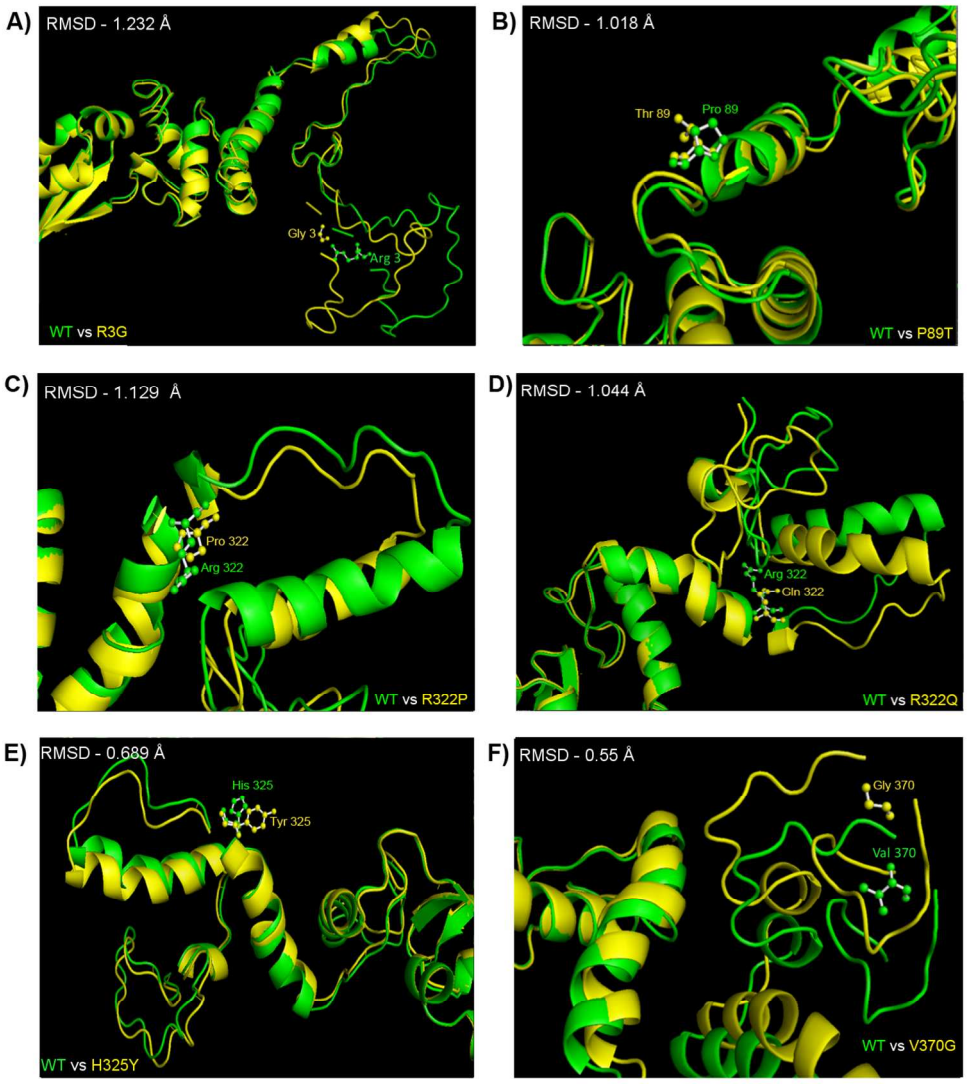

Figure 6: Structural alignment of podocin and its mutants: Each figure (A-F) represents structural alignment of wild type podocin model with a mutant. The ball and stick representation in each figure (A: WT vs R3G; B: WT vs P89T; C: WT vs R322P; D: WT vs R322Q; E: WT vs H325Y; F: WT vs V370G) shows the amino acid that was mutated. RMSD for each mutant was calculated using Pymol.
365x412mm (96 x 96 DPI)

Table 1: Prediction of intrinsically unstructured regions (IUR) in podocin by Critical Assessment of Protein Structure Prediction (CASP) validated servers.

Server	Residues	Remarks
Genesilico Meta disorder server	1 - 121, 337 - 383	Meta-disorder 3D
	1 - 100, 338 - 383	Meta-disorder MD2
	1 - 98, 347 - 383	Meta-disorder MD
	1 - 97, 347 - 383	Meta-disorder
PONDR	1 - 96, 230 - 242, 268 - 315, 322 - 325, 357 - 383	-
PSI - Pred	1- 7 (BD), 8-61, 63 - 69 (BD), 77-89 (BD), 90-94, 353 - 356 (BD), 357 - 362, 363 (BD), 364 - 366, 367 - 377 (BD), 378 - 383	(BD) – Binding Domains
PredictProtein	1 - 94, 351 - 383	-
DIS - PROT	1 - 94, 216 - 226, 277 - 323, 336 - 383	-
DIS - EMBL	1 - 78, 86 - 99, 141 - 163, 168- 183, 225 - 276, 327 - 383	Disordered by Loops/ Coils Definition
	1 - 76, 84 - 100, 359 - 383	Disordered by HOT-Loops
	1 - 59, 86 - 96, 355 - 383	Remark - 465
IUPred	1 - 70, 78 - 88, 355 - 383	-
SPINE - D	1-125, 344 - 383	-
MFDp	1 - 96, 340 - 383	MFDp
	1 - 101, 279 - 290, 296, 298 - 316, 342 - 383	DISOclust
	1 - 98, 308 - 310, 342 - 383	DISOPred
	1 - 70, 78 - 88, 350 - 383	IUPredL
	1 - 88, 352 - 383	IUPredS
OnDCRF	1-100, 347-383	Predicting ordered and disordered regions using conditional random fields

Table 2: Prediction of intrinsically unstructured regions (IURs) in stomatin family proteins.

Proteins	Residues
Erythrocyte band 7 integral membrane protein	1-25, 203-241, 278-288
Hflc	1-3, 134 -197, 228-273, 318 - 334
Hflk	1-79, 210-223, 249-287, 330-419
Protein UNC-1	1-25, 196-239, 279-285
Stomatin - 1	1-36, 202-249, 286-330
Stomatin - 2	1-117, 297-340, 353-374
Stomatin – 3	1-7, 193-209, 259-267
Stomatin-like protein – 1	1-49, 221-292, 395-398
Stomatin-like protein - 3	1-22, 197-242, 259-263, 274-288

Table 3: Distribution of backbone dihedral angles of identified mutants

S. No	Mutants	Ramachandran plot analysis			
		Most favoured	Additionally allowed	Generously allowed	Disallowed
1.	R3G	86.6%	10.3%	2.4%	0.6%
2.	P89T	85.8%	11.5%	1.8%	0.9%
3.	R322Q	87.0%	10.6%	1.8%	0.6%
4.	R322P	85.4%	11.9%	2.1%	0.6%
5.	H325Y	86.1%	10.9%	2.4%	0.6%
6.	V370G	86.0%	11.2%	1.8%	0.9%